

AN EXTENSION OF THE FUNCTIONAL ITO FORMULA UNDER A FAMILY OF NON-DOMINATED MEASURES

HARALD OBERHAUSER

ABSTRACT. Motivated by questions arising in financial mathematics, Dupire [12] introduced a notion of smoothness for functionals of paths (different from the usual Fréchet–Gâteaux derivatives) and arrived at a generalization of Itô’s formula applicable to functionals which have a pathwise continuous dependence on the trajectories of the underlying process. We study nonlinear functionals which do not have such pathwise continuity and further work simultaneously under the family of continuous semimartingale measures on path-space. We do this without introducing a second component, as carried out by Cont–Fournie [9, 8] but by using old work of Bichteler [4] which allows to keep a pathwise picture even for complex functionals.

1. INTRODUCTION

It is an understatement to say that Itô’s stochastic calculus is a useful tool for the modeling of systems that evolve under the influence of randomness. An essential part of this calculus is Itô’s formula which, given a semimartingale $X = (X_t)_{t \geq 0}$ and a sufficiently smooth function f , shows that the stochastic process $f(t, X_t)$ is also a semimartingale and moreover, provides an explicit (Bichteler–Dellacherie) decomposition of the process $f(t, X_t)$ into a sum of a stochastic integral against X and a process of finite variation. Of course, the class of processes of the form $f(t, X_t)$ is only a small subset of the stochastic processes that are adapted to the filtration generated by X , $\sigma(X)$, and often it is necessary to derive similar statements for this larger class of processes. One area where such questions arise is financial engineering: traders aim to understand the dynamics of the price process of a (highly path-dependent) contingent claim with respect to the underlying asset modeled by the stochastic process X . Motivated by these questions, Dupire [12] showed that if X is a standard real-valued Brownian motion, it is possible to extend Itô’s formula to a nontrivial subset in the class of real-valued processes $F = (F_t)_{t \geq 0}$ which are adapted to $\sigma(X)$ and gave the formula

$$(1.1) \quad F_t - F_0 = \int_0^t \left(\partial_0 F_r + \frac{1}{2} \Delta F_r \right) dr + \int_0^t \nabla F_r dX$$

where $\partial_0 F, \nabla F$ and ΔF are again $\sigma(X)$ -adapted processes to which we refer as the causal derivatives of F (following M. Fliess¹; in [12, 9, 8] these are called functional time and space derivatives or also horizontal and vertical derivatives). For the special case $F_t = f(t, X_t)$, these causal derivatives coincide with the usual time and space derivatives of f and the above reduces to the standard Itô-formula applied to $f(t, X_t)$. However, the class of processes considered by Dupire is limited to those that depend continuously² on the trajectories of X (e.g. $F = \int_0^\cdot f(r, X_r) dr$ but not $F = \int_0^\cdot f(r, X_r, [X]_r) dr$ etc.). While such a path-by-path continuity is of course guaranteed by differentiability in the case of the classic Itô-formula, $F_t = f(t, X_t)$, it is not implied for general, $\sigma(X)$ -adapted processes F by causal differentiability; indeed, such a pathwise continuity would be a very strong restriction in the class of $\sigma(X)$ -adapted processes. This restriction was subsequently addressed by Cont–Fournie who consider general semimartingales X and add a second process (e.g. $[X]$ or its weak derivative; cf. [9, 8] and the excellent thesis of D. Fournie) to express F as a functional which depends pathwise continuously on the trajectories of X and this second process which allows them to arrive at a generalization of Dupire’s functional Itô-formula to a larger class of $\sigma(X)$ -adapted processes. This article is inspired by all these strong and beautiful results and shares the same goal of extending the class of functionals of X for which (1.1) holds, however, we explore a different approach. We do not introduce a second process in

Key words and phrases. Semimartingale decomposition, pathwise stochastic Integration, functional Ito formula.

¹We use the term causal since essentially the same operators can be found in the literature on nonlinear system control (i.e. deterministic, bounded variation paths), c.f. [18, Section IIb], where they are known as causal derivatives.

²in (not precisely but) essentially the uniform topology, cf. Example 13 and [9]; in fact, much of this article is concerned with avoiding the use of this metric.

addition to the continuous semimartingale X but nevertheless arrive at a generalisation of (1.1) to a large class of functionals which includes adapted processes with a complex dependence on the past like the quadratic variation of X , stochastic integrals, Doléans–Dade exponentials, compositions thereof, etc. To do this we have to overcome some obstacles:

- firstly, the pathwise nature of the causal derivatives involves operations on sets of paths which are null-sets,
- secondly, many interesting processes are constructed by a probabilistic argument, hence are often only uniquely defined as equivalence classes of indistinguishable processes but not pathwise unique,
- thirdly, that many of these processes are not robust under approximations of X in uniform norm.

Especially note the clash of the first point (pathwise considerations matter) and the second point (only equivalence classes under a fixed probability measure matter for many applications in stochastic analysis). We briefly elaborate on this and sketch our approach: Throughout we work on the canonical path-space $(\Omega, \mathcal{F}^0, (\mathcal{F}_t^0))$ of \mathbb{R}^d -valued (possibly defective, cf. Section 1.1) càdlàg processes with $X_t(\omega) = \omega(t)$ denoting the coordinate process. In Section 2 we recall Dupire’s causal derivatives. They are a family of maps $\mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ which describe the sensitivity of a given $F : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ with respect to perturbations of the coordinate process X at running time. If we denote with \mathcal{M}_c^{semi} the probability measures under which the coordinate process X is a semimartingale with (\mathbb{P} -a.s) continuous trajectories then we can consider the standard completion $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$ of $(\Omega, \mathcal{F}^0, \mathcal{F}_t^0)$ (the reason why we work on the space of càdlàg paths Ω is the definition of the causal space derivative). Now assume we are given an adapted, real-valued process F on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$, $\mathbb{P} \in \mathcal{M}_c^{semi}$ then even if we have a version of this F at hand which has causal derivatives then it is, as pointed out above, not always justified to make strong assumptions on the pathwise regularity of the map $(t, \omega) \mapsto F_t(\omega)$ (already for fixed t ; e.g. if F is any version of the stochastic integral, quadratic variation etc.). Therefore we introduce in Section 3 the class of functionals $C^{1,2}$, i.e. a subset of the maps $\mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$, and show in Section 4 that the functional Itô formula extends to $C^{1,2}$; Theorem 22 on page 8 reads

Theorem. *Let $F : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$, $F \in C^{1,2}$. Then $\forall \mathbb{P} \in \mathcal{M}_c^{semi}$, F is a continuous semimartingale on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$ and*

$$F_t - F_0 = \int_0^t \partial_0 F_r dr + \sum_{i=1}^d \int_0^t \partial_i F_r dX_r^i + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \partial_{ij} F_r d[X^i, X^j]_r \quad \mathbb{P} - a.s.$$

Of course, it remains to show that an interesting class of processes can be expressed as a functional in $C^{1,2}$: for the class of processes studied by Dupire (Example 13 on page 7) this follows immediately. In Section 5 we do exactly this for the stochastic integral $\int Y_- dX$ and the quadratic variation process $[X^i, X^j]$ using the pathwise Itô-integral due to Bichteler and Karandikar, for example we show in Section 5

Proposition. *For every $i, j \in \{1, \dots, d\}$ there exists a map $B^{ij} : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ such that $B^{ij} \in C^{1,2}$ and B^{ij} is on every $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$, $\mathbb{P} \in \mathcal{M}_c^{semi}$, an adapted process, indistinguishable from the quadratic variation process $[X^i, X^j]$ (constructed under \mathbb{P}).*

Note that this is a non-trivial statement linking the map/functional B^{ij} with the quadratic variation process $[X^i, X^j]$, the latter being only uniquely defined as an equivalence class of indistinguishable processes on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$ for every fixed $\mathbb{P} \in \mathcal{M}_c^{semi}$. The difficulty is to choose a representant in this class which is an element of $C^{1,2}$ and additionally also a representant for the quadratic variation constructed under any other measure in \mathcal{M}_c^{semi} , i.e. to find an aggregator³ which is also differentiable — if we would restrict ourselves to a subset $\mathcal{P} \subset \mathcal{M}_c^{semi}$ which is dominated by a single measure, i.e. all elements in \mathcal{P} are absolutely continuous wrt to this measure, then at least the existence of the aggregator would be trivial though not necessarily its differentiability. A similar result holds for the Itô-integral $\int Y_- \cdot dX$ and as a consequence of the results in Section 5 we get

³Given a family \mathcal{P} of probability measures and a family of processes $\{F^\mathbb{P} : \mathbb{P} \in \mathcal{P}, F^\mathbb{P} \text{ is a meas. process on } (\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P})\}$ we call $F : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ an aggregator of this family if $F = F^\mathbb{P}$ \mathbb{P} -a.s. $\forall \mathbb{P} \in \mathcal{P}$ (see [33] for a slightly different definition).

Proposition. *Let $\mu, \sigma : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ be sufficiently regular (as in Theorem 33 resp. Example 43). Then there exists a map $I : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ such that $I \in C^{1,2}$ and I is on every $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P}), \mathbb{P} \in \mathcal{M}_c^{semi}$, an adapted process, indistinguishable from the process (constructed under \mathbb{P})*

$$(1.2) \quad \int_0^\cdot \mu_r dr + \int_0^\cdot \sigma_{r-} dX_r.$$

Moreover, $\partial_0 F = \mu$, $(\partial_i F)_{i=1,\dots,d} = (\sigma_-^i)_{i=1,\dots,d}$ and $(\partial_{ij} F)_{i,j=1,\dots,d} = 0$ \mathbb{P} -a.s.

If we denote with \mathcal{P}_{ac} the subset of \mathcal{M}_c^{semi} under which $t \mapsto [X]_t$ is absolutely continuous wrt to Lebesgue measure then we can sum up the above in the language of quasi-sure analysis [11, 33]: $F = \overline{F}$ \mathcal{P}_{ac} -quasi-sure for some $\overline{F} \in C^{1,2}$ iff F is of the form (1.2) \mathcal{P}_{ac} -quasi-sure⁴ (under suitable assumptions on the coefficients μ, σ); further, the class $C^{1,2}$ contains aggregators with complex pathdependence. We emphasize that we do not extend the causal derivatives by closure of operators on equivalence classes of indistinguishable processes to cover Itô-processes which leads to strong results under one fixed measure \mathbb{P} but develop a pathwise picture applicable simultaneously to all $\mathbb{P} \in \mathcal{M}_c^{semi}$. The approach to consider a stochastic calculus under a family of non-dominated probability measures is motivated by recent developments in probability theory like Denis' and Martini's uncertain volatility model and quasi-sure analysis via aggregation [33, 11] as well as Peng's G -expectation [29, 31]. Of course, plugging in above functionals B and I into the functional Itô formula leads to trivial identities but the point is that compositions with smooth functions or functionals are again in $C^{1,2}$ (e.g. see the examples in Section 5 or [20] for more applications in finance). Further, the only requirement for a functional $F : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ for being in $C^{1,2}$ is besides its causal differentiability (and some minor regularity assumptions) the convergence of F (and its derivatives) under finite-dimensional approximations, $F(\omega^n) \rightarrow_n F(\omega)$ uniformly on compacts in \mathbb{P} -probability, $\forall \mathbb{P} \in \mathcal{M}_c^{semi}$, where $(\omega^n)_n$ is a piecewise constant approximation to ω , which — in view of the usual approximation results in Itô-calculus — seems to be quite a weak and natural assumption (recall classic approximations in Itô-calculus for say the stochastic integral or quadratic variation which more or less by construction hold in probability or even Wong–Zakai type⁵ results for highly complex, pathdependent functionals are in principle included).

1.1. Notation. Denote with ζ an isolated point added to \mathbb{R}^d which plays the role of a cemetery and denote with \mathbb{R}_+ the set $[0, \infty)$. We work on the canonical path-space for sub-stochastic (i.e. possibly defective) càdlàg processes denoted with Ω , i.e. Ω is the set of paths $\omega : \mathbb{R}_+ \rightarrow \mathbb{R}^d \cup \{\zeta\}$ with lifetime

$$\zeta(\omega) = \inf \{t \geq 0 : \omega(t) = \zeta\}$$

which are càdlàg and stay at the cemetery ζ after their lifetime⁶ $\zeta(\omega)$. Denote the coordinate process $(X_t)_{t \in \mathbb{R}_+}$,

$$X_t(\omega) := \omega(t)$$

as well as

$$X_{t-}(\omega) := \lim_{s \uparrow t} X_s(\omega), \quad \Delta_t X(\omega) := X_t(\omega) - X_{t-}(\omega).$$

Further, we introduce a σ -field and filtration on Ω ,

$$\mathcal{F}^0 := \sigma(X_s, 0 \leq s) \text{ and } \mathcal{F}_t^0 := \sigma(X_s, 0 \leq s \leq t).$$

Note that Ω can be endowed with the Skorohod topology in which case the Borel σ -algebra equals \mathcal{F}^0 and with this topology Ω is a Polish space (cf. [24, Chapter VI, Theorem 1.4] or [3, Chapter 0]; however we are not making use of this Polish structure in this article). To speak about predictable processes we have to define \mathcal{F}_{0-} and X_{0-} which we simply define as $\mathcal{F}_{0-} = \{\emptyset, \Omega\}$, $X_{0-} = \zeta$; further denote the Borel σ -field of \mathbb{R}_+ with $\mathcal{B}_{\mathbb{R}_+}$ and the Lebesgue measure on \mathbb{R}_+ with λ . As usual we call any collection of random variables on Ω indexed by time $t \in \mathbb{R}_+$ a stochastic process and a stochastic process (F_t) is said to be measurable if $(t, \omega) \mapsto F_t(\omega)$ is measurable on $\mathbb{R}_+ \times \Omega$ with respect to

⁴An event holds \mathcal{P} -quasi-sure if it holds \mathbb{P} -a.s. $\forall \mathbb{P} \in \mathcal{P}$ where \mathcal{P} is a subset of \mathcal{M}_c^{semi} (cf. [11, 16]).

⁵For obvious reasons, Wong–Zakai results are usually formulated for the convergence along piecewise linear paths (not piecewise constant) hence giving a Stratonovich- (not an Itô-) calculus.

⁶With abuse of notation we use ζ for the isolated point as well as the map $\zeta : \Omega \rightarrow [0, \infty]$. Since $\{\omega : \zeta(\omega) < t\} = \bigcup_{r < t, r \in \mathbb{Q}} \{\omega : X_r(\omega) = \zeta\} \in \mathcal{F}_t^0$ it follows that $\zeta(\cdot)$ is a random variable. In fact the cemetery ζ is not essential for our arguments but it allows for intuitive characterizations of predictable processes (Proposition 1) which is useful for some proofs.

$\mathcal{B}_{\mathbb{R}_+} \times \mathcal{F}^0$; similarly (F_t) is progressively measurable if for each $t \in \mathbb{R}_+$ the map $[0, t] \times \Omega \rightarrow \mathbb{R}$, $(s, \omega) \mapsto F_s(\omega)$ is $(\mathcal{B}_{[0, t]} \times \mathcal{F}_t^0)$ -measurable and a process (F_t) is said to be adapted to (\mathcal{F}_t^0) if for each $t \in \mathbb{R}_+$, F_t is \mathcal{F}_t^0 -measurable. The optional σ -field \mathcal{O}^0 on $\mathbb{R}_+ \times \Omega$ is generated by the real-valued càdlàg processes adapted to $(\mathcal{F}_t^0)_{t \geq 0}$ and the predictable σ -field \mathcal{P}^0 on $\mathbb{R}_+ \times \Omega$ is generated by the real-valued processes adapted to⁷ $(\mathcal{F}_{t-}^0)_{t \geq 0}$ with càg (left-continuous) paths on $(0, \infty)$, cf. [10, p121-IV] for further properties of these σ -algebras. Given $(t, \omega) \in \mathbb{R}_+ \times \Omega$ and $r \in \mathbb{R}^d \cup \{\zeta\}$, we denote $\omega^{t, r} \in \Omega$ the càdlàg path which coincides until time t with ω but has a jump at time t in direction $r \in \mathbb{R}^d \cup \{\zeta\}$ and stays constant after time t , viz.

$$\omega^{t, r}(s) = \omega(s \wedge t) + r 1_{s \geq t}, \quad s \in \mathbb{R}_+$$

(with the convention $a + \zeta = \zeta$ for any $a \in \mathbb{R}^d \cup \{\zeta\}$); further denote with $\omega_{\wedge t}$ the path ω stopped at time t , viz.

$$\omega_{\wedge t}(s) = \omega(s \wedge t), \quad s \in \mathbb{R}_+.$$

These two perturbations of the path ω play a central role in this article and lead to the definition of a causal time and space derivative of a process in Section 2. However, at this point we only recall that such perturbations arise naturally when working on the canonical path-space $(\Omega, \mathcal{F}^0, (\mathcal{F}_t^0))$ and lead to an intuitive characterization of predictable and optional processes.

Proposition 1 (Dellacherie–Meyer [10, p147-IV]).

- (1) The predictable σ -field \mathcal{P}^0 on $\mathbb{R}_+ \times \Omega$ is generated by the two maps

$$(t, \omega) \mapsto t \text{ and } (t, \omega) \mapsto \omega^{t, \zeta}$$

and the optional σ -field \mathcal{O}^0 on $\mathbb{R}_+ \times \Omega$ is generated by the two maps

$$(t, \omega) \mapsto t \text{ and } (t, \omega) \mapsto \omega_{\wedge t}.$$

- (2) A measurable process F is predictable (wrt to \mathcal{P}^0) iff

$$F_t(\omega) = F_t(\omega^{t, \zeta}) \quad \forall (t, \omega) \in \mathbb{R}_+ \times \Omega$$

and a measurable process F is optional (wrt to \mathcal{O}^0) iff

$$F_t(\omega) = F_t(\omega_{\wedge t}) \quad \forall (t, \omega) \in \mathbb{R}_+ \times \Omega.$$

We say that a sequence $(\pi(n))_n$ of partitions of \mathbb{R}_+ (i.e. $\pi(n) = \{t_k^n\}_{k \geq 0}$ where $t_k^n \in \mathbb{R}_+$, $t_k^n \leq t_{k+1}^n$ and $t_0^n = 0$) converges to the identity if

$$\sup_{k \geq 0} |t_k^n - t_{k-1}^n| \rightarrow 0 \text{ and } \sup_{k \geq 0} t_k^n \rightarrow +\infty \text{ as } n \rightarrow \infty.$$

Denote for a given $t \in \mathbb{R}_+$ with $t^{\pi(n)}$ resp. $t_{\pi(n)}$ the nearest elements of the partition $\pi(n)$ to the right resp. left of t with the convention $t \in [t_{\pi(n)}, t^{\pi(n)}]$; similarly we use ${}^{\pi(n)}t$ resp. ${}_{\pi(n)}t$ with the convention $t \in ({}_{\pi(n)}t, {}^{\pi(n)}t]$. Further define for $\omega \in \Omega$ the path $\omega^{\pi(n)} \in \Omega$ as the piecewise constant approximation⁸ of ω along $\pi(n)$

$$(1.3) \quad \omega^{\pi(n)}(t) = \omega(0) + \sum_{k: {}_{\pi(n)}t_{k-1} \in \pi(n)} (\omega(t_k^n) - \omega(t_{k-1}^n)) 1_{t \geq t_k^n}.$$

Given a probability measure \mathbb{P} denote with $(\Omega, \mathcal{F}^{\mathbb{P}}, \mathcal{F}_t^{\mathbb{P}}, \mathbb{P})$ the usual augmentation (i.e. right-continuous and complete) of $(\Omega, \mathcal{F}^0, \mathcal{F}_t^0, \mathbb{P})$. Denote with \mathcal{M}_c^{semi} the set of probability measures \mathbb{P} under which the coordinate process X is a semimartingale with \mathbb{P} -a.s. continuous trajectories and similarly with \mathcal{M}_c resp. \mathcal{M}_c^{loc} the space of measures under which X is a martingale resp. locale martingale with \mathbb{P} -a.s. trajectories. Above objects and terminology (measurable, the optional and predictable σ -fields, etc.) can be defined with respect to the completed σ -algebra $\mathcal{F}^{\mathbb{P}}$ resp. $(\mathcal{F}_t^{\mathbb{P}})$ instead of \mathcal{F}^0 and (\mathcal{F}_t^0) , e.g. giving rise to the predictable σ -field $\mathcal{P}^{\mathbb{P}}$, the optional σ -field $\mathcal{O}^{\mathbb{P}}$ (instead of $\mathcal{O}^0, \mathcal{P}^0$) etc. Recall that two processes F and G are said to be indistinguishable/a version of each other on $(\Omega, \mathcal{F}^{\mathbb{P}}, \mathcal{F}_t^{\mathbb{P}}, \mathbb{P})$ if $\mathbb{P}(F_t = G_t \quad \forall t \in \mathbb{R}_+) = 1$; since parts of this article require us to argue pathwise (i.e. with a fixed

⁷this is equivalent except at $t = 0$ to being adapted to $(\mathcal{F}_t^0)_{t \geq 0}$. To quote Dellacherie–Meyer [10, p121-IV], “in all considerations on the predictable σ -field, time 0 plays the devil’s role”.

⁸With the convention $\zeta + r = \zeta$ for $r \in \mathbb{R}^d \cup \{\zeta\}$.

version rather than modulo indistinguishability) we decided to be rather pedantic about spelling out quantifiers for all statements.

2. CAUSAL FUNCTIONALS AND CAUSAL DERIVATIVES

Definition 2. We say that a map $F : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ is a causal functional if

$$F_t(\omega_{\wedge t}) = F_t(\omega) \quad \forall (t, \omega) \in \mathbb{R}_+ \times \Omega.$$

Remark 3. We do not require measurability of causal functionals with respect to the uncompleted σ -algebra $(\mathcal{B}_{\mathbb{R}_{\geq 0}} \times \mathcal{F}^0)$; in other words not every causal functional F is a process on $(\Omega, \mathcal{F}^0, \mathcal{F}_t^0)$ but if F is causal and $(\mathcal{B}_{\mathbb{R}_{\geq 0}} \times \mathcal{F}^0)$ -measurable then F is an \mathcal{O}^0 -optional process on the canonical path-space $(\Omega, \mathcal{F}^0, \mathcal{F}_t^0)$ (by Proposition 1, see also Remark 14). Of course, every \mathcal{O}^0 -optional process is a causal functional.

Following Dupire [12] (see also earlier work of Fliess [18, 17] in the context of bounded variation paths from which took the term *causal*), we define the time and space derivatives of causal functionals.

Definition 4. Let $F : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ be a causal functional. If $\forall (t, \omega) \in \mathbb{R}_+ \times \Omega$, $t < \zeta(\omega)$ the map

$$\mathbb{R}^d \ni r \mapsto F_t(\omega^{t,r}) \in \mathbb{R}$$

is continuously differentiable at $r = 0$ then we denote with $\nabla F_t(\omega) = (\partial_1 F_t(\omega), \dots, \partial_d F_t(\omega))$ its Jacobian and set $\nabla F_t(\omega) = 0$ for (t, ω) with $t \geq \zeta(\omega)$. We then say that F has $\nabla F : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}^d$ as causal space derivative. Similarly, we define the matrix-valued second causal space derivative $\Delta F = (\partial_i \partial_j F)_{i,j=1}^d$ as the Hessian.

Definition 5. Let $F : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ be a causal functional. If $\forall (t, \omega) \in \mathbb{R}_+ \times \Omega$, $t < \zeta(\omega)$ the map

$$\mathbb{R}_+ \ni r \mapsto F_{t+r}(\omega_{\wedge t}) \in \mathbb{R}$$

is continuous and has a right-derivative at $r = 0$ then we denote with $\partial_0 F_t(\omega)$ this right derivative and set $\partial_0 F_t(\omega) = 0$ if $t \geq \zeta(\omega)$. If additionally $r \mapsto \partial_0 F_r(\omega)$ is Riemann integrable then we say that F has the map $\partial_0 F : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ as causal time derivative.

Remark 6. These derivatives can be seen as natural extension of the usual time and space derivatives of functions, e.g. if $F_t(\omega) = f(t, X(t, \omega))$ and $f \in C^{1,1}(\mathbb{R}_+ \times \mathbb{R}, \mathbb{R})$ then $\partial_0 F_t(\omega) = \frac{df}{dt}(t, X(t, \omega))$ and $\nabla F_t(\omega) = \frac{df}{dx}(t, X(t, \omega))$ (if $t < \zeta(\omega)$). Further, it is easy to verify the analogues of the standard rules of differentiation, i.e. chain and product rules, etc.

Remark 7. If ω and $\tilde{\omega}$ coincide on $[0, t]$ then $\nabla F_r(\omega) = \nabla F_r(\tilde{\omega}) \quad \forall r \leq t$ and the same statement is true for ΔF and $\partial_0 F$, i.e. the causal derivatives of a causal functional are again causal functionals.

Remark 8. These *causal derivatives are defined pathwise, hence do not respect null-sets on $(\Omega, \mathcal{F}^{\mathbb{P}}, \mathcal{F}_t^{\mathbb{P}}, \mathbb{P})$* for a given $\mathbb{P} \in \mathcal{M}_c^{semi}$ (unlike the Fréchet derivative in directions of the Cameron–Martin space if \mathbb{P} is Gaussian, as used by Bismut, Malliavin et al.). It may happen that two processes F and G are indistinguishable on $(\Omega, \mathcal{F}^{\mathbb{P}}, \mathcal{F}_t^{\mathbb{P}}, \mathbb{P})$, both have causal time and space derivatives but apriori there is no reason why these causal derivatives should actually be indistinguishable from each other⁹. To sum up, versions matter and we need an additional regularity assumption.

Note that throughout this section we avoided the language of probability/measure theory and above remark shows how things can go wrong — in fact above remark might disturb our reader because it implies that we have to drop the elegant, usual approach of stochastic analysis to regard two processes as equal if they are indistinguishable under the measure \mathbb{P} . Instead we work pathwise. We hope to reconcile her in Section 4 where it turns out that under an additional assumption on F (namely regularity as introduced in the section below *and* causal differentiability) the examples in Remark 8 are to a large extent excluded (Proposition 29). In the mean time we simply ask for her patience and mention that situations in stochastic analysis in which indistinguishability is not a sufficient criteria are not too uncommon: prominent examples include Clark’s robustness problem in nonlinear filtering [7], quasi-sure analysis via aggregation [11, 33], pathwise expansions [6] or the recent interest in pathwise delta-hedging arguments where in all these cases the existence of a “good pathwise version” is important.

⁹e.g. let $F_t(\omega) = 0$, $\tilde{F}_t(\omega) = 1_{\Delta_t X(\omega) \neq 0}$ and $\overline{F}_t(\omega) = c \Delta_t X(\omega)$ for some $c \in \mathbb{R}$. Then F, \tilde{F} and \overline{F} are indistinguishable on $(\Omega, \mathcal{F}^{\mathbb{P}}, \mathcal{F}_t^{\mathbb{P}}, \mathbb{P}) \quad \forall \mathbb{P} \in \mathcal{M}_c^{semi}$, F has causal time and space derivatives (all equal 0), however \tilde{F} does not even have a causal space derivative and \overline{F} has a space derivative equal to the arbitrary chosen c !

3. REGULAR AND DIFFERENTIABLE CAUSAL FUNCTIONALS

We have to cope with the problem that functionals on an infinite-dimensional (path-)space can be causal differentiable but not continuous with respect to uniform topology. A simple solution is to additionally require such a continuity, however, since we are interested in functionals of unbounded variation paths this is indeed a very strong assumption. This section introduces a class of functionals, regular enough to prove a functional Itô formula but still rich enough to include aggregators of the stochastic integral, quadratic variation, Doléans–Dade exponential etc.

3.1. Regular, causal functionals. The proof of the functional Itô formula requires to understand the behaviour of causal functionals under pathwise approximations of the coordinate process X . We briefly recall that pathwise approximations of functionals of unbounded variation paths are more subtle than the case of bounded variation paths, already under a fixed measure $\mathbb{P} \in \mathcal{M}_c^{semi}$.

Example 9. Fix $\mathbb{P} \in \mathcal{M}_c^{semi}$ and consider the stochastic process on $(\Omega, \mathcal{F}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$ defined as

$$F_t(\omega) = \left(\int_0^t X^1 dX^2 \right)(\omega) - \left(\int_0^t X^2 dX^1 \right)(\omega)$$

(any version of the stochastic integral on the right hand side), i.e. Lévy’s area process. To the best of our knowledge, the first explicit example of an approximation of the process X which leads to a correction term is due to McShane [27] for the case when X is a Brownian motion on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$, that is for an arbitrary $c \in \mathbb{R}$ there exists for \mathbb{P} -a.e. ω a sequence $(\omega^n)_n \subset \Omega$ s.t. $|X(\omega^n) - X(\omega)|_{[0,t],\infty} \rightarrow_n 0$ but the process $F(\omega^n)$ converges to the process $(F_t + c.t)_t$ \mathbb{P} -a.s. — we do not spell out the details (but refer to the excellent presentation in [27, 23, Chapter VI] in the context of Wong–Zakai approximations) though we briefly recall the main idea: consider the dyadic partitions and interpolate ω between two points not linear, but by choosing between two interpolating functions to construct ω^n . The key is to make this choice dependent on the sign of (an approximation of) the Lévy area surpassed in this time interval¹⁰. This highly-oscillatory perturbation is not canceled in the limit as the mesh size tends to 0 and picked up in an additive correction term of the iterated integrals of X . Also note that such a behaviour is not a question of Fisk–Stratonovich vs. Itô integration (for the Lévy-area F it does not matter which integration is used since the quadratic variation brackets cancel). Sussman (cf. [35, 22]) even shows that for arbitrary $N \in \mathbb{N}$ one can find approximations such that the first N iterated (Stratonovich) integrals converge to the iterated integrals of X but lead to a correction term on the $(N + 1)$ -th level of iterated integrals.

A classic strategy to cope with such phenomena is to restrict attention to a class of “natural” finite-dimensional approximations, typically piecewise constant or piecewise linear approximations (depending on an Itô or a Stratonovich approach), e.g. to prove Wong–Zakai theorems [37, 38, 23, Chapter VI], large deviation results à la Varadhan, Freidlin–Wentzell, Azencott [36, 21, 2], or Stroock–Varadhan type support theorems [34] etc. In view of an Itô-calculus and that the causal derivatives are defined by piecewise constant perturbations resp. perturbations by jumps, it is not surprising that for us the “natural class” of approximations are the piecewise constant approximations of the coordinate process $X(\omega) = \omega$ (for the definition below recall that ω^π is the piecewise constant approximation of ω along π ; as in (1.3)).

Definition 10. We say a causal functional $F : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ is a weakly regular, càdlàg functional if for every $\mathbb{P} \in \mathcal{M}_c^{semi}$ there exists a sequence of partitions $\pi = (\pi(n))_n$, converging to identity such that

- (1) F is an $(\mathcal{F}_t^\mathbb{P})$ -adapted càdlàg process on¹¹ $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$,
- (2) the $(\mathcal{F}_t^\mathbb{P})$ -adapted, measurable process $(t, \omega) \mapsto F_t^n(\omega) := F_t(\omega^{\pi(n)})$ has left limits (for every n \mathbb{P} -a.s.) and $F_t^n \rightarrow F_-$ as $n \rightarrow \infty$ uniformly on compacts in \mathbb{P} -probability (henceforth ucp¹²).

If point (2) holds for every sequence of partitions $\pi = (\pi(n))_n$ converging to identity and every $\mathbb{P} \in \mathcal{M}_c^{semi}$ then we simply say that F is a regular, càdlàg functional. Similarly, we say F is a

¹⁰McShane’s and Sussman’s approximations are not adapted but, like the usual piecewise linear approximation, they can be simply shifted one interval back in time to make them adapted.

¹¹i.e. $t \mapsto F_t$ is \mathbb{P} -a.s. càdlàg

¹²That is, $\forall T \in \mathbb{R}_+$ we have $\sup_{r \in [0, T]} |F_{r-}^n - F_{r-}| \rightarrow_{n \rightarrow \infty} 0$ in probability on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$.

regular càglàd (resp. continuous) functional if in point (1) we replace the word càdlàg by càglàd (resp. continuous).

Remark 11. Instead of point (2) we could simply demand that F_t^n is a càdlàg (or càglàd) process and $F^n \rightarrow_n F$ ucp. This would already include a large class of functionals, however, the above, more general formulation allows to include functionals F which only have left limits along the approximation F^n .

Remark 12. It is easy to write down a functional which is only weakly regular but not regular. However, Section 5 shows that even processes with complex path-dependence have a “canonical version” which is regular, i.e. a regular aggregator which is also causally differentiable.

The important part of Definition 10 is point (2) since point (1) is by Proposition 1 already satisfied whenever the causal functional is a measurable process with càdlàg trajectories on the canonical probability space $(\Omega, \mathcal{F}^0, \mathcal{F}_t^0)$. An interesting subclass of functionals which are regular are the ones treated by Dupire [12].

Example 13. Let $F : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ be an \mathcal{O}^0 -optional process on $(\Omega, \mathcal{F}^0, \mathcal{F}_t^0)$ and as a map uniformly continuous on compacts of \mathbb{R}_+ with respect to the pseudo-metric on $\mathbb{R}_+ \times \Omega \setminus \{\omega : \zeta(\omega) < \infty\}$

$$d_{Dupire}((t, \omega), (\bar{t}, \bar{\omega})) = |t - \bar{t}| + |\omega_{\wedge t} - \bar{\omega}_{\wedge \bar{t}}|_{\infty; [0, t \vee \bar{t}]}.$$

This is the class of processes studied by Dupire [12]. Note that $\mathcal{O}^0 \subset \mathcal{O}^{\mathbb{P}}$ and the continuity with respect to $d_{Dupire}(\cdot, \cdot)$ guarantees that F is causal and also the convergence of $F_t(\omega^{\pi(n)})$ as required by Definition 10 (even pathwise!). Hence, every such F is a causal, regular continuous functional. While this is quite a small class of processes in view of the usual processes of interest in stochastic analysis, it already allows for interesting examples (especially in mathematical finance where the process F can model the price evolution of a path-dependent option).

Remark 14. In Definition 10 of a regular functional we do not require F to be a measurable process on $(\Omega, \mathcal{F}^0, \mathcal{F}_t^0)$ (by Proposition 1 only measurability of F is needed to make a causal functional an \mathcal{O}^0 -optional process) but only measurability on the completed probability spaces $(\Omega, \mathcal{F}^{\mathbb{P}}, (\mathcal{F}_t^{\mathbb{P}}), \mathbb{P})$. This is a minor technical point but this generality turns out to be useful when dealing with causal functionals with very complex pathdependence where measurability with respect to the uncompleted σ -algebra \mathcal{F}^0 resp. filtration (\mathcal{F}_t^0) can be hard to establish (e.g. in the pathwise Bichteler integral where “stopping/hitting times” appear etc., see Section 5).

3.2. The class $C^{1,2}$ of causal functionals. In this section, we introduce the class $C^{1,2}$ of causal functionals which are regular and have regular, causal derivatives. Before we do this we need another definition which will be only needed for the causal derivatives of a functional.

3.3. Causal continuity in time and space.

Definition 15. We say a causal functional $F : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ is causally continuous in space, if $\forall R > 0$, $\forall t > 0$ there exists a function $\rho_{t,R} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ which is non-decreasing and $\lim_{x \searrow 0} \rho_{t,R}(x) = 0$ such that

$$\sup_{r \in [0, t]} \left| F_r(\omega^{r, \Delta_r}) - F_r(\omega^{r, \tilde{\Delta}_r}) \right| = \rho_{t,R} \left(\left| \Delta - \tilde{\Delta} \right|_{\infty; [0, t]} \right)$$

holds $\forall \omega, \Delta, \tilde{\Delta} \in \Omega$ with $|\omega|_{\infty; [0, t]}, |\Delta|_{\infty; [0, t]}, |\tilde{\Delta}|_{\infty; [0, t]} \leq R$.

Similarly, we give the corresponding definition in the time variable.

Definition 16. We say a causal functional $F : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ is causally continuous in time, if $\forall R > 0$, $\forall t > 0$ there exists a function $\rho_{t,R} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ which is non-decreasing and $\lim_{x \searrow 0} \rho_{t,R}(x) = 0$ such that

$$\sup_{r \in [0, t]} |F_{r+\Delta_r}(\omega_{\wedge r}) - F_r(\omega)| = \rho_{t,R} \left(|\Delta|_{\infty; [0, t]} + \sup_{r \in [0, t]} |\omega_r - \omega_{r-}| \right)$$

holds $\forall \omega \in \Omega, \forall \Delta \in C([0, t], \mathbb{R}_+)$ with $|\omega|_{\infty; [0, t]}, |\Delta|_{\infty; [0, t]} \leq R$.

Remark 17. To motivate above definitions note that they are immediately fulfilled for $F_t(\omega) = f(t, X_t(\omega))$ whenever $f \in C([0, T] \times \mathbb{R}^n, \mathbb{R})$ (since f is uniformly continuous on compacts, hence has a modulus). Intuitively the continuity in space means that F does react to jumps in the underlying in proportion to the jump size and the continuity in time implies that jumps of F are only due to jumps in the underlying (not due to progression of time alone).

Remark 18. If F is continuous with respect to d_{Dupire} (Example 13) then F is causally continuous in time and space and the latter are much weaker requirements for functionals than pathwise continuity in the sense of Example 13 since we only require continuity with respect to jumps resp. stopping the path at running time, not continuity with respect to perturbations of the whole past of the path (e.g. causal continuity in time and space hold for the pathwise version of the stochastic integral or quadratic variation in Section 5).

3.4. The class $C^{1,2}$. We now introduce the class of functionals we are interested in.

Definition 19. We denote with $C^{1,2}$ the set of all causal functionals $F : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ which fulfill

- (1) F is a regular continuous functional which has one causal time and two causal space derivatives,
- (2) $\partial_0 F, \nabla F, \Delta F$ are regular càglàd or regular càdlàg functionals,
- (3) $\nabla F, \Delta F$ are causally continuous in time and space.

Remark 20. For functions $f(t, X_t)$ continuity is guaranteed by differentiability, however for causal functionals the questions of causal differentiability and continuity wrt to the underlying have to be treated separately. This is not surprising since we work with nonlinear functionals on an infinite dimensional (path-)space.

Remark 21. We immediately see that $\{(t, \omega) \mapsto f(t, \omega(t)) : f \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R})\} \subset C^{1,2}$, i.e. as our notation reveals, we think of $C^{1,2}$ as the analogue for general $\sigma(X)$ -adapted processes of the class of processes $f(t, X_t)$ for $f \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$. Recall Example 13 which shows that differentiable processes with pathwise continuous dependence on the underlying are included in $C^{1,2}$. But most important for us, the results of Section 5 below show that $C^{1,2}$ includes functionals which depend on the past of $X(\omega) = \omega$ in a more complex way like stochastic integrals, quadratic variation, Doléans–Dade exponentials, compositions thereof, etc.

We have arrived at a point comparable to the moment in a basic analysis course where differentiability of functions has been introduced and it remains to show that many basic functions are differentiable and to calculate their derivative. In other words, we still have to show that $C^{1,2}$ contains aggregators of interesting processes (besides the examples covered by Dupire [12] resp. Example 13). This trite observation is the motivation for Section 5. However, before that, we show in the section below that the functional Itô-formula extends to the class of $C^{1,2}$ functionals.

4. THE FUNCTIONAL ITÔ FORMULA FOR $C^{1,2}$ -FUNCTIONALS

We are now in position to prove an extension of the functional Itô-formula to $C^{1,2}$. The basic idea of the proof is similar to a standard proof of the usual Itô-formula, which is to carry out a Taylor expansion of order 2, and — following Dupire — the role of the usual derivatives is taken by the causal time and space derivatives.

Theorem 22. *If $F \in C^{1,2}$ then F is a continuous semimartingale on $(\Omega, \mathcal{F}^{\mathbb{P}}, \mathcal{F}_t^{\mathbb{P}}, \mathbb{P}) \forall \mathbb{P} \in \mathcal{M}_c^{semi}$ and*

$$(4.1) \quad F_t - F_0 = \int_0^t \partial_0 F_r dr + \sum_{i=1}^d \int_0^t \partial_i F_r dX_r^i + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \partial_{ij} F_r d[X^i, X^j]_r \quad \mathbb{P} - a.s.$$

Proof. Fix $t > 0$ and a sequence $(\pi(n))_n$ of partitions $\pi(n) = (t_k^n)_k$ converging to the identity. For brevity write ω^n for the piecewise constant càdlàg approximation $\omega^{\pi(n)}$ of ω along $\pi(n)$ and $r^n, {}^n r$ resp. $r_{n,n}, {}^n r$ for the nearest element $\pi(n)$ to the right resp. left of r (with our usual convention $r \in [r_n, r^n)$ and $r \in ({}_n r, {}^n r]$), further we use the notation $X_{s,t}(\omega) := \omega_{s,t} := \omega(t) - \omega(s)$, $s, t \in \mathbb{R}_+$. Throughout assume that n is big enough s.t. $t < \sup_k t_k^n$. For every $\omega \in \Omega$ with $\zeta(\omega) > t$ write

$$(4.2) \quad \begin{aligned} F_{t^n}(\omega^n) - F_0(\omega^n) &= F_{t^n}(\omega_{\wedge t^n}^n) - F_0(\omega_{\wedge 0}^n) \\ &= \sum_k F_{t_k^n}(\omega_{\wedge t_k^n}^n) - F_{t_{k-1}^n}(\omega_{\wedge t_{k-1}^n}^n) \end{aligned}$$

(causality of F implies $F_{t^n}(\omega^n) = F_{t^n}(\omega_{\wedge t^n}^n)$ and $F_0(\omega^n) = F_0(\omega_{\wedge 0}^n)$) where the sum \sum_k is taken over all k s.t. $\pi(n) \ni t_{k-1}^n, t_k^n \leq t^n$ and split the terms on the right hand side into

$$F_{t_k^n}(\omega_{\wedge t_k^n}^n) - F_{t_{k-1}^n}(\omega_{\wedge t_{k-1}^n}^n) = \underbrace{F_{t_k^n}(\omega_{\wedge t_k^n}^n) - F_{t_k^n}(\omega_{\wedge t_{k-1}^n}^n)}_{=:S_k^n(\omega)} + \underbrace{F_{t_k^n}(\omega_{\wedge t_{k-1}^n}^n) - F_{t_{k-1}^n}(\omega_{\wedge t_{k-1}^n}^n)}_{=:T_k^n(\omega)}.$$

The intuition is that S_k^n measures the space increment at time t_k^n and T_k^n measures the time decay on the interval $[t_{k-1}^n, t_k^n]$. In step 1 and 2 below we show that for every $\mathbb{P} \in \mathcal{M}_c^{semi}$ we have on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$ the following convergence in \mathbb{P} -probability

$$(4.3) \quad \sum_k S_k^n \xrightarrow{n \rightarrow \infty} \int_0^t \nabla F_{r-} \cdot dX_r + \frac{1}{2} \int_0^t Tr[\Delta F_{r-} \cdot d[X]_r],$$

$$(4.4) \quad \sum_k T_k^n \xrightarrow{n \rightarrow \infty} \int_0^t \partial_0 F_r dr.$$

Further, since F is a regular functional the process $(t, \omega) \mapsto F_{t-}(\omega^n)$ converges ucp to F_- and since $t \mapsto F_t$ is \mathbb{P} -a.s. continuous (4.2) converges in probability to $F_t - F_0$. This together with (4.3) and (4.4) implies (4.1) for fixed t . The continuity of F then allows to remove the time dependence of the null-set.

Step 1: $\sum_k S_k^n \xrightarrow{\mathbb{P}} \int_0^t \nabla F_- dX + \frac{1}{2} \int_0^t Tr[\Delta F_- d[X]]$. Fix $\omega \in \Omega$ with $\zeta(\omega) = \infty$ and note that $\omega_{\wedge t_k^n}^n = (\omega_{\wedge t_{k-1}^n}^n)^{t_k^n, \omega_{t_{k-1}^n}^n, t_k^n}$. Taylor's theorem applied to the function

$$\mathbb{R}^d \ni \bullet \mapsto F_{t_k^n} \left((\omega_{\wedge t_{k-1}^n}^n)^{t_k^n, \bullet} \right) \in \mathbb{R}$$

guarantees the existence of a $\theta_k^n(\omega) \in [0, 1]$ s.t.

$$\begin{aligned} F_{t_k^n}(\omega_{\wedge t_k^n}^n) - F_{t_{k-1}^n}(\omega_{\wedge t_{k-1}^n}^n) &= \sum_i \partial_i F_{t_k^n}(\omega_{\wedge t_{k-1}^n}^n) \omega_{t_{k-1}^n, t_k^n}^i \\ &\quad + \frac{1}{2} \sum_{i,j} \partial_{ij} F_{t_k^n} \left((\omega_{\wedge t_{k-1}^n}^n)^{t_k^n, \theta_k^n(\omega) \omega_{t_{k-1}^n}^n, t_k^n} \right) \omega_{t_{k-1}^n, t_k^n}^i \omega_{t_{k-1}^n, t_k^n}^j. \end{aligned}$$

Hence

$$(4.5) \quad F_{t_k^n}(\omega_{\wedge t_k^n}^n) - F_{t_{k-1}^n}(\omega_{\wedge t_{k-1}^n}^n) = \sum_i \partial_i F_{t_k^n}(\omega_{\wedge t_{k-1}^n}^n) \omega_{t_{k-1}^n, t_k^n}^i + \frac{1}{2} \sum_{i,j} \partial_{ij} F_{t_k^n}(\omega_{\wedge t_{k-1}^n}^n) \omega_{t_{k-1}^n, t_k^n}^i \omega_{t_{k-1}^n, t_k^n}^j + r_{n,k}(\omega, \omega_{t_{k-1}^n}^n, t_k^n)$$

where

$$(4.6) \quad r_{n,k}(\omega, \Delta) = \frac{1}{2} \sum_{i,j} \left[\partial_{ij} F_{t_k^n} \left((\omega_{\wedge t_{k-1}^n}^n)^{t_k^n, \theta_k^n(\omega) \Delta} \right) - \partial_{ij} F_{t_k^n}(\omega_{\wedge t_{k-1}^n}^n) \right] \Delta^i \Delta^j \quad \forall \Delta \in \mathbb{R}^d.$$

By the causal continuity in space of $\Delta F = (\partial_{ij} F)_{i,j=1,\dots,d}$ we can estimate (using that $\omega_{\wedge t_{k-1}^n}^n = (\omega^n)^{t_k^n, -\omega_{t_{k-1}^n}^n, t_k^n}, (\omega_{\wedge t_{k-1}^n}^n)^{t_k^n, \theta_k^n(\omega) \Delta} = (\omega^n)^{t_k^n, -\omega_{t_{k-1}^n}^n, t_k^n + \theta_k^n(\omega) \Delta}, |\omega_{t_{k-1}^n, t_k^n}^n| \leq 2|\omega|, |\omega^n| \leq 2|\omega|$)

$$\begin{aligned} &\left| \partial_{ij} F_{t_k^n} \left((\omega_{\wedge t_{k-1}^n}^n)^{t_k^n, \theta_k^n(\omega) \omega_{t_{k-1}^n}^n, t_k^n} \right) - \partial_{ij} F_{t_k^n}(\omega_{\wedge t_{k-1}^n}^n) \right| \\ &= \left| \partial_{ij} F_{t_k^n} \left((\omega^n)^{t_k^n, -\omega_{t_{k-1}^n}^n, t_k^n + \theta_k^n(\omega) \omega_{t_{k-1}^n}^n, t_k^n} \right) - \partial_{ij} F_{t_k^n} \left((\omega^n)^{t_k^n, -\omega_{t_{k-1}^n}^n, t_k^n} \right) \right| \\ (4.7) \quad &\leq \rho_{t,2|\omega|} \left(\left| \theta_k^n(\omega) \omega_{t_{k-1}^n, t_k^n}^n \right| \right) \leq \rho_{t,2|\omega|} \left(\left| \omega_{t_{k-1}^n, t_k^n}^n \right| \right) \end{aligned}$$

with $\rho_{t,2|\omega|}(0+) = 0$. Taking \sum_k in (4.5) yields

$$(4.8) \quad \begin{aligned} \sum_k S_k^n(\omega) &= \sum_k \nabla F_{t_k^n} \left(\omega_{\wedge t_{k-1}^n}^n \right) \cdot X_{t_{k-1}^n, t_k^n}^n(\omega) \\ &\quad + \frac{1}{2} \sum_k Tr \left[X_{t_{k-1}^n, t_k^n}^T(\omega) \cdot \Delta F_{t_k^n} \left(\omega_{\wedge t_{k-1}^n}^n \right) \cdot X_{t_{k-1}^n, t_k^n}^n(\omega) \right] \\ &\quad + R_n(\omega) \end{aligned}$$

with $R_n(\omega) := \sum_k r_{n,k} \left(\omega, \omega_{t_{k-1}^n, t_k^n}^n \right)$ and by (4.6) and (4.7) we arrive at

$$\begin{aligned} |R_n(\omega)| &\leq \frac{1}{2} \sum_k \sum_{i,j} \left| \partial_{ij} F_{t_k^n} \left(\left(\omega_{\wedge t_{k-1}^n}^n \right)^{t_k^n, \theta_k^n(\omega) \omega_{t_{k-1}^n, t_k^n}^n} \right) - \partial_{ij} F_{t_k^n} \left(\omega_{\wedge t_{k-1}^n}^n \right) \right| \left| \omega_{t_{k-1}^n, t_k^n}^i \omega_{t_{k-1}^n, t_k^n}^j \right| \\ &\leq \frac{1}{2} \sup_k \rho_{t,2|\omega|} \left(\left| \omega_{t_{k-1}^n, t_k^n}^n \right| \right) \sum_k \sum_{i,j} \left| X_{t_{k-1}^n, t_k^n}^i(\omega) \right| \left| X_{t_{k-1}^n, t_k^n}^j(\omega) \right| \\ &\leq \frac{1}{2} \rho_{t,2|\omega|} \left(\sup_k \left| \omega_{t_{k-1}^n, t_k^n}^n \right| \right) \sum_{i,j} \sum_k \left(\left| X_{t_{k-1}^n, t_k^n}^i(\omega) \right|^2 + \left| X_{t_{k-1}^n, t_k^n}^j(\omega) \right|^2 \right). \end{aligned}$$

Note that $\sum_k \left(\left| X_{t_{k-1}^n, t_k^n}^i(\omega) \right|^2 + \left| X_{t_{k-1}^n, t_k^n}^j(\omega) \right|^2 \right) \rightarrow_n [X^i]_t + [X^j]_t$ in probability and obviously for \mathbb{P} -a.e. ω we have

$$\rho_{t,2|\omega|} \left(\sup_k \left| \omega_{t_{k-1}^n, t_k^n}^n \right| \right) \rightarrow_n 0$$

hence $|R_n(\cdot)| \rightarrow_n 0$ in probability. It remains to show convergence of the first two sums on the rhs in (4.8). Therefore we rewrite $\sum_k \nabla F_{t_k^n} \left(\omega_{\wedge t_{k-1}^n}^n \right) \cdot X_{t_{k-1}^n, t_k^n}^n(\omega)$ by adding and subtracting $\nabla F_{t_{k-1}^n} \left(\omega_{\wedge t_{k-1}^n}^n \right)$ and see that it is sufficient to prove

$$\begin{aligned} \sum_k \nabla F_{t_{k-1}^n} \left(\omega_{\wedge t_{k-1}^n}^n \right) \cdot X_{t_{k-1}^n, t_k^n}^n(\omega) &\rightarrow_{n \rightarrow \infty}^{\mathbb{P}} \left(\int_0^t \nabla F_- \cdot dX \right)(\omega) \\ \sum_k \left(\nabla F_{t_k^n} \left(\omega_{\wedge t_{k-1}^n}^n \right) - \nabla F_{t_{k-1}^n} \left(\omega_{\wedge t_{k-1}^n}^n \right) \right) \cdot X_{t_{k-1}^n, t_k^n}^n(\omega) &\rightarrow_{n \rightarrow \infty}^{\mathbb{P}} 0. \end{aligned}$$

For the first sum note that $\nabla F_{t_{k-1}^n} \left(\omega_{\wedge t_{k-1}^n}^n \right) = \nabla F_{t_{k-1}^n}(\omega^n)$. For brevity we introduce the processes $I^n, (t, \omega) \mapsto I_t^n(\omega) := \nabla F_t(\omega^n)$ and $I_t(\omega) := \nabla F_t(\omega)$ and note that $I_-^n \rightarrow_n I_-$ in \mathbb{P} -probability (since ∇F is regular by assumption). For \mathbb{P} -a.e. ω we have

$$\sum_k \nabla F_{t_{k-1}^n}(\omega^n) \cdot X_{t_{k-1}^n, t_k^n}^n(\omega) = \left(\int I_{nr}^n dX_r \right)(\omega) = \left(\int I_{nr} dX_r \right)(\omega) + \left(\int (I_{nr}^n - I_{nr}) dX_r \right)(\omega)$$

(recall the notation $r \in (nr, {}^n r]$). Note that $(I_{nr})_{r \in \mathbb{R}_+}$ is a piecewise constant càglàd process (therefore a simple integrand), hence $\int I_{nr} dX_r$ converges ucp to $\int_0^t I_- dX \equiv \int_0^t \nabla F_- dX$; since I_-^n converges ucp to $I_- \equiv \nabla F_-$ we get that $\int (I_{nr}^n - I_{nr}) dX$ converges ucp to 0. It remains to estimate

$$(4.9) \quad \sum_k \left(\nabla F_{t_k^n} \left(\omega_{\wedge t_{k-1}^n}^n \right) - \nabla F_{t_{k-1}^n} \left(\omega_{\wedge t_{k-1}^n}^n \right) \right) \cdot X_{t_{k-1}^n, t_k^n}^n(\omega).$$

Therefore note that $\left(\nabla F_{t_k^n} \left(\cdot_{\wedge t_{k-1}^n}^n \right) - \nabla F_{t_{k-1}^n} \left(\cdot_{\wedge t_{k-1}^n}^n \right) \right) \in \mathcal{F}_{t_{k-1}^n}^{\mathbb{P}}$, hence above sum can again be seen as the stochastic integral of a simple integrand and causal continuity in time implies

$$\sup_k \left| \nabla F_{t_k^n} \left(\omega_{\wedge t_{k-1}^n}^n \right) - \nabla F_{t_{k-1}^n} \left(\omega_{\wedge t_{k-1}^n}^n \right) \right| \leq \rho_{t,2|\omega|} \left(mesh(\pi_n) + \sup_k \left(\omega_{t_{k-1}^n}^n - \omega_{t_{k-2}^n}^n \right) \right) \rightarrow_n 0 \text{ for } \mathbb{P}\text{-a.e. } \omega$$

which is sufficient to conclude that (4.9) converges to 0 in probability. To sum up, we have shown that the sum $\sum_k \nabla F_{t_{k+1}^n} \left(\omega_{\wedge t_{k-1}^n}^n \right) \cdot X_{t_{k-1}^n, t_k^n}^n(\omega)$ appearing in (4.8) converges ucp to $\int_0^t \nabla F_- dX$. The very same arguments imply the convergence

$$\frac{1}{2} \sum_k Tr \left[X_{t_{k-1}^n, t_k^n}^T(\omega) \cdot \Delta F_{t_k^n} \left(\omega_{\wedge t_{k-1}^n}^n \right) \cdot X_{t_{k-1}^n, t_k^n}^n(\omega) \right] \rightarrow_{n \rightarrow \infty}^{\mathbb{P}} \int_0^t Tr [\Delta F_- \cdot d[X]].$$

Step 2: $\sum_k T_k^n \xrightarrow{\mathbb{P}} \int_0^t \partial_0 F_r dr$. Causal time differentiability of F implies that

$$\mathbb{R}_+ \ni \bullet \mapsto F_{t_{k-1}^n + \bullet} \left(\omega_{\wedge t_{k-1}^n}^n \right)$$

is continuous and has a right-derivative on¹³ \mathbb{R}_+ which is additionally Riemann integrable. This is enough for a modification of the fundamental theorem of Riemann integration to apply (see [5]):

$$F_{t_{k-1}^n + (t_k^n - t_{k-1}^n)} \left(\omega_{\wedge t_{k-1}^n}^n \right) - F_{t_{k-1}^n} \left(\omega_{\wedge t_{k-1}^n}^n \right) = \int_0^{t_k^n - t_{k-1}^n} \partial_0 F_{t_{k-1}^n + r} \left(\omega_{\wedge t_{k-1}^n}^n \right) dr.$$

Taking the sum \sum_k over all k s.t. $\pi(n) \ni t_{k-1}^n, t_k^n \leq t^n$ gives

$$\sum_k T_k^n = \int_0^{t^n} \partial_0 F_r \left(\omega_{\wedge r_n}^n \right) dr.$$

Since $\partial_0 F$ is causal, $\partial_0 F_r \left(\omega_{\wedge r_n}^n \right) = \partial_0 F_r \left(\omega^n \right)$ for $r \notin \pi(n)$ (which is a Lebesgue null-set of $[0, t]$), we have shown

$$\sum_k T_k^n = \int_0^{t^n} \partial_0 F_r \left(\omega^n \right) dr.$$

By assumption the integrand converges ucp on $(\Omega, \mathcal{F}^{\mathbb{P}}, \mathcal{F}_t^{\mathbb{P}}, \mathbb{P})$ to $\partial_0 F$, hence we have shown the convergence

$$\sum_k T_k^n \xrightarrow{\mathbb{P}}_{n \rightarrow \infty} \int_0^t \partial_0 F_r dr.$$

□

Remark 23. Theorem 22 says that $F \in C^{1,2}$ implies that F is an Itô-process with regular coefficients. In Section 5 we show (Theorem 33 and Example 43) the converse, i.e. every process of the form $\int_0^t \mu_r dr + \int_0^t \sigma_r dX_r$ has (provided μ, σ are regular enough) an aggregator in $C^{1,2}$. Theorem 22 immediately gives the explicit Bichteler–Dellacherie–Meyer decomposition of F under every $\mathbb{P} \in \mathcal{M}_c^{loc} \subset \mathcal{M}_c^{semi}$.

Remark 24. Dupire showed above formula for the case when $(t, \omega) \mapsto F(t, \omega)$ is pathwise continuous in the (pseudo-)metric of Example 13 and \mathbb{P} is the Wiener measure. Cont and Fournie [9, 8] give a mathematical precise argument and two proofs: a pathwise, analytic proof in [8] and a probabilistic proof in [9]. Further, both apply to functionals which depend pathwise continuously on ω and a second component.

Remark 25. Motivated by the case of no pathdependence, i.e. $F_t(\omega) = f(t)$ one could argue that a natural criteria for a fundamental theorem of calculus is absolute continuity instead of right-differentiability. In fact, alternatively to Definition 5, one could say that F has a causal time derivative if $r \mapsto F_r(\omega_{\wedge t})$ is absolutely continuous and above proof then applies modulo obvious modifications.

Remark 26. We restricted attention to continuous semimartingale measures \mathcal{M}_c^{semi} (i.e. $t \mapsto X_t$ is a.s. continuous) but under extra assumptions on the regularity of F one can also in the case when X is a càdlàg semimartingale subtract the small jumps before steps 1 and 2 (see also [8]). However, the assumptions on F get more technical and in view the applications given in the sections below we decided to stick to the case \mathcal{M}_c^{semi} since our focus lies on the lack of continuity $\omega \mapsto F(\omega)$. In the same spirit one could ask for a Tanaka formula etc.

The proof of Theorem 22 requires only regularity of F and its derivatives along one sequence of partitions converging to identity and with respect to a fixed measure $\mathbb{P} \in \mathcal{M}_c^{semi}$, i.e. we have shown a more general statement: denote with $C_w^{1,2}$ the class of functionals which fulfill Definition 19 when we replace regular with the condition that for every $\mathbb{P} \in \mathcal{M}_c^{semi}$ there exists a sequence of partitions $\pi = (\pi(n))$ converging to identity such that F and its derivatives are weakly-regular along π .

¹³Not only at $\bullet = 0$ since $F_{(t_{k-1}^n + r) + \bullet} \left(\omega_{\wedge t_{k-1}^n}^n \right) = F_{(t_{k-1}^n + r) + \bullet} \left(\left(\omega_{\wedge t_{k-1}^n}^n \right)_{\wedge (t_{k-1}^n + r)} \right) \forall r \in \mathbb{R}_+$, hence $F_{t_{k-1}^n + r + \bullet} \left(\omega_{\wedge t_{k-1}^n}^n \right) - F_{t_{k-1}^n + r} \left(\omega_{\wedge t_{k-1}^n}^n \right) = \partial_0 F_{t_{k-1}^n + r} \left(\omega_{\wedge t_{k-1}^n}^n \right) \bullet + o(\bullet)$ as $\bullet \searrow 0$.

Corollary 27. *If $F \in C_w^{1,2}$ then F is a continuous semimartingale on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$ and*

$$F_t - F_0 = \int_0^t \partial_0 F_r dr + \sum_{i=1}^d \int_0^t \partial_i F_r dX_r^i + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \partial_{ij} F_r d[X^i, X^j]_r \quad \mathbb{P} - a.s.$$

Applied with $F_t(\omega) = f(t, X_t(\omega))$ all the above reduces to the standard Itô formula and motivated by this one can ask for generalizations of the Feynman–Kac formula or the classic link between (sub-)harmonic functions and (sub-)martingales. Indeed, all this was already done by Dupire [12] and one can follow the same arguments the only difference being that we now cover a larger class of aggregators – though we still have to convince the reader of this, i.e. that $C^{1,2}$ includes interesting functionals besides the functionals of Example 13. Before that we briefly come back to the non-uniqueness of the causal derivatives seen as processes as pointed out in Remark 8.

4.1. Uniqueness of causal heat operator and space derivative acting on processes. Due to the pathwise nature of the causal derivatives, it may happen that two different functionals give rise to processes that are indistinguishable, both are differentiable but these derivatives are no longer indistinguishable processes.

Example 28. Let $d = 1$. In Section 5 we introduce a causal functional, denoted $[X]^{BK} : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$, s.t. $[X]^{BK} \in C^{1,2}$ and under every $\mathbb{P} \in \mathcal{M}_c^{semi}$, $[X]^{BK}$ is indistinguishable on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$ from the quadratic variation process of X on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$. Consider the two causal functionals, both are in $C^{1,2}$,

$$(t, \omega) \mapsto [X]_t^{BK}(\omega) \quad \text{and} \quad (t, \omega) \mapsto t$$

which are indistinguishable processes if X is a Brownian motion under \mathbb{P} however their causal derivatives are distinguishable under \mathbb{P} (e.g. $\partial_0(t) = 1$ but $\partial_0([X]^{BK}) = 0$ \mathbb{P} -a.s., cf. Corollary 39).

However, the situation is not too bad: in above example the first spatial derivative ∇ and $\partial_0 + \frac{1}{2}Tr[\Delta \cdot d[X]]$ applied to these functionals are again indistinguishable processes on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$. Theorem 22 gives us the simple explanation that $\partial_0 F$ and $\frac{1}{2}\Delta F$ both replicate the behaviour of F on the time-scale of t as opposed to ∇F which replicates the behaviour of F on the \sqrt{t} -scale, hence we can only expect uniqueness of the “causal heat operator” $\partial_0 F + \frac{1}{2}Tr[\Delta F d[X]]$ and ∇F .

Proposition 29. *Let $F, G \in C^{1,2}$, $\mathbb{P} \in \mathcal{M}_c^{loc}$ and assume $[X]_\cdot = \int_0^\cdot \sigma^2 dr$ under \mathbb{P} for some continuous, $\mathcal{F}_t^\mathbb{P}$ -adapted process σ . Further assume F and G are indistinguishable on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$ i.e.*

$$F_\cdot = G_\cdot \quad \mathbb{P} - a.s$$

Then

$$(4.10) \quad \partial_0 F + \frac{1}{2}Tr[\Delta F \cdot \sigma^2] = \partial_0 G + \frac{1}{2}Tr[\Delta G \cdot \sigma^2] \quad \mathbb{P} - a.s.$$

and

$$(4.11) \quad \int_0^\cdot \nabla F \cdot dX = \int_0^\cdot \nabla G \cdot dX \quad \mathbb{P} - a.s.$$

Proof. As in [9], this is a simple consequence of the uniqueness of the semimartingale decomposition: from Theorem 22 it follows that $H_t(\omega) := F_t(\omega) - G_t(\omega)$ is a continuous semimartingale on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$ but every continuous semimartingale is a special semimartingale, that is the decomposition $H = M + A$ into a local martingale M and a process A of locally bounded variation is unique up to \mathbb{P} -indistinguishability which already implies (4.11). Since H is indistinguishable from the process 0 this implies

$$A_\cdot = \int_0^\cdot \left(\partial_0 H_r + \frac{1}{2}Tr[\Delta H_r \cdot \sigma_r^2] \right) dr$$

is indistinguishable from 0, hence we have that \mathbb{P} -a.s. $\partial_0 H_t + \frac{1}{2}Tr[\Delta H_t \cdot \sigma_t^2] = 0$ Lebesgue a.e. which implies (4.10). \square

Remark 30. Similar path-dependent heat operators appear already in seminal work of Kusuoka–Stroock [26] using Malliavin calculus techniques on abstract Wiener spaces. However, no Feynman–Kac formula was given and the abstract Wiener space setting is of course essential for the results in [26]. Further, we mention the interesting approach of Ahn [1] on semimartingale representations via Fréchet derivatives.

Remark 31. Under stronger assumptions on \mathbb{P} (e.g. when X is a non-degenerate Brownian martingale) one can also deduce the indistinguishability of ∇F and ∇G since then $\int \nabla F \cdot dX = \int \nabla G \cdot dX$ already implies indistinguishability of ∇F and ∇G . This is essential to define an extension of the causal derivatives ∇ and $\partial_0 + \frac{1}{2}Tr[\Delta \cdot d[X]]$ which are defined pathwise to operators on (equivalence classes of) Itô-processes via closure available under a fixed measure as in Malliavin calculus cf. [9, Section 5] and Remark 34. We note that in classic work of Kusuoka–Stroock [26] a similar phenomenon appears where the Fréchet derivative D and only the “Malliavin heat operator” $\mathcal{A} = \frac{\partial}{\partial t} + \frac{1}{2}T_H D^2$ (not only $\frac{\partial}{\partial t}$ or $\frac{1}{2}T_H D^2$) are extended by continuity and closure on abstract Wiener spaces (cf. [26, p11, Remark 1.17ff]).

5. THE ITÔ-INTEGRAL AND QUADRATIC VARIATION PROCESS AS CAUSAL FUNCTIONALS

In this section we show that one can construct a pathwise version of the stochastic integral of a càdlàg process against the coordinate process X which is an element of $C^{1,2}$, i.e. a regular, differentiable causal functional and similarly for the quadratic variation. To achieve this we make use of the path-wise Itô-integral and quadratic variation as introduced by Bichteler in 1981 using adapted Riemann sums and subsequent simplifications by Karandikar [4, 25].

Example 32. To appreciate Bichteler’s pathwise integral let us briefly recall why pathwise constructions are non-trivial on the simpler example of the quadratic variation process in the one-dimensional case ($d = 1$). Once a probability measure $\mathbb{P} \in \mathcal{M}_c^{semi}$ and a sequence of partitions $(\pi(n))_n$ is fixed, standard results guarantee the existence of a process $[X]$ on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$ (for brevity let $d = 1$) such that

$$X_0^2 + \sum_{k:t_{k-1}^n \in \pi(n)} \left| X_{t_k^n \wedge t} - X_{t_{k-1}^n \wedge t} \right|^2 \rightarrow_n [X]_t \quad \text{uniformly on compacts in } \mathbb{P}\text{-probability.}$$

Hence, a candidate for a pathwise version (which is also causally differentiable) of the quadratic variation process is the map

$$(t, \omega) \mapsto F_t(\omega) = \begin{cases} |\omega(0)|^2 + \lim_{n \rightarrow \infty} \sum_{k:t_{k-1}^n \in \pi(n)} \left| \omega(t_k^n \wedge t) - \omega(t_{k-1}^n \wedge t) \right|^2 & , \text{ if this limit exists} \\ 0 & , \text{ otherwise.} \end{cases}$$

Indeed, if $\mathbb{P} \in \mathcal{M}_c^{semi}$ is such that the coordinate process X is a Brownian motion and $(\pi(n))_n$ are dyadics then above F is a process on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$ indistinguishable from $[X]$. However, this is not true under every $\mathbb{P} \in \mathcal{M}_c^{semi}$ or any sequence of partitions $(\pi(n))_n$ converging to the identity — in fact, $\sum_k \left| \omega(t_k^n \wedge t) - \omega(t_{k-1}^n \wedge t) \right|^2$ converges in general only in \mathbb{P} -probability. (Of course, for every given $\mathbb{P} \in \mathcal{M}_c^{semi}$ one can always find a fast enough vanishing sub-sequence of $(\pi(n))_n$ to ensure convergence \mathbb{P} -a.s.) The same problem appears, afortiori, for the stochastic integral where path-regularity of the integrand has to be considered as well, etc.

The main result of this section is Theorem 33 below which shows that the causal space derivative acts as the inverse of a pathwise stochastic integral.

Theorem 33. *Assume Z is a real-valued càdlàg process on $(\Omega, \mathcal{F}_t^0, \mathcal{F}^0)$ which is càdlàg regular. Then for every $i \in \{1, \dots, d\}$ there exists a regular càdlàg functional*

$$J_Z^i : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$$

such that

- (1) J_Z^i has causal derivatives $\partial_0 J_Z^i, \nabla J_Z^i, \Delta J_Z^i$ and these are again regular functionals,
- (2) J_Z^i is under every $\mathbb{P} \in \mathcal{M}_c^{semi}$ indistinguishable from the Itô-integral $\int_0^\cdot Z_- dX^i$ on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$, i.e.

$$J_Z^i(\cdot, \omega) = \left(\int_0^\cdot Z_- dX^i \right)(\omega) \quad \text{for } \mathbb{P}\text{-a.e. } \omega.$$

Moreover $\forall (t, \omega) \in \mathbb{R}_+ \times \Omega, t < \zeta(\omega)$,

$$\begin{aligned} \partial_0 J_Z^i(t, \omega) &= 0, \\ \nabla J_Z^i(t, \omega) &= (0, \dots, 0, Z_{t-}(\omega), 0, \dots, 0)^T \quad (\text{only } i^{\text{th}} \text{ component non-zero}), \\ \Delta J_Z^i(t, \omega) &= 0 \end{aligned}$$

and if Z_- is causally continuous in time then $J_Z^i \in C^{1,2}$. We also use the notation $\int_0^\cdot Z_-(\omega) d^{BK} X^i(\omega)$ for the process $J_Z^i(\omega)$ (d^{BK} stands for Bichteler–Karandikar), resp. in the multidimensional case $Z = (Z^i)_{i=1}^d$ we write $\int_0^\cdot Z_-(\omega) \cdot d^{BK} X(\omega)$ for $\sum_{i=1}^d J_Z^i(\omega)$.

Remark 34. In [9, Section 5] ∇ is extended under a fixed measure \mathbb{P} to an operator ∇_X acting on square-integrable (standard) stochastic Itô-integrals which can be identified via the Clark–Hausmann–Ocone Theorem as the predictable projection of the Malliavin derivative. We refer the reader to [9, Section 5], but remark that Theorem 33 shows that the causal derivative ∇ acts as the *pathwise inverse* to an aggregator of the Itô-integral, e.g. in the case of a Wiener functional $\xi \in L^2(\mathcal{F}_T^\mathbb{P})$ that has the representation $\xi = \mathbb{E}_\mathbb{P}[\xi] + \int_0^T Z_- \cdot dX$ with Z as in Theorem 33 then ξ has a representation as ξ_T^0 , $\xi^0 \in C^{1,2}$, given as $\xi_t^0(\omega) := \mathbb{E}_\mathbb{P}[\xi^0] + \int_0^t \nabla \xi_-^0(\omega) \cdot d^{BK} X(\omega)$; more interestingly this is not restricted to Wiener functionals provided of course the existence of such a representation $\xi = m + \int_0^T Z_- \cdot dX$. The considerable price is that already in the Brownian context one has to assume some regularity of Z but martingale representation does not guarantee any such regularity. However it might be a reasonable assumption for applications (e.g. implementable Delta-hedging strategies, integrands of the form $Z_t = v(t, X_t, [X]_t^{BK})$ etc; note also that using the non-pathwise L^2 -extension ∇_X leads to not completely trivial issues about consistent approximations by the pathwise operator ∇ acting on cylinder functions).

The process J_Z^i in Theorem 33 is constructed via the pathwise Bichteler–Karandikar integral which we briefly recall.

Theorem 35 (Bichteler [4], Karandikar [25, Theorem 2 and 3]). *There exists a map*

$$I : D(\mathbb{R}_+, \mathbb{R}) \times D(\mathbb{R}_+, \mathbb{R}) \rightarrow D(\mathbb{R}_+, \mathbb{R})$$

s.t. for any filtered probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t), \tilde{\mathbb{P}})$ satisfying the usual conditions, any real-valued semimartingale \tilde{X} and any adapted, real-valued, càdlàg process \tilde{Z} on $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t), \tilde{\mathbb{P}})$ we have

$$I(\tilde{Z}(\omega), \tilde{X}(\omega))(\cdot) = \left(\int_0^\cdot \tilde{Z}_- d\tilde{X} \right)(\omega) \quad \tilde{\mathbb{P}}\text{-a.s.}$$

Remark 36. The set of full $\tilde{\mathbb{P}}$ -measure on which the equality holds depends of course on the processes \tilde{Z} and \tilde{X} .

Remark 37. Another well-known result on pathwise integration is Föllmer’s Itô-integral “sans probabilités” [19]. However, this allows only to give a pathwise version when the integrand is of the type $\int \nabla f(X_s) \cdot dX_s$, i.e. a non-pathdependent integrand which is in gradient form (the latter is a strong restriction in multi-dimensions $d > 1$, cf. Example 9 for a process which is not of this type). The non-pathdependence can be relaxed, see [8], however the gradient form is essential for Föllmer’s argument. Let us also mention the approach via capacity as in Denis and Martini [11] to construct an integral for quasi-sure integrands and Peng’s approach via PDEs [28] which lead to a much less explicit integral but on the other hand apply to less regular integrands than the Bichteler–Karandikar integral.

Remark 38. The map I is constructed as follows: for $x, z \in D(\mathbb{R}_+, \mathbb{R})$ and $n \in \mathbb{N}$ define the sequence $(\tau_i^n)_{i \geq 0} \subset \mathbb{R}_+ \cup \{+\infty\}$ recursively as¹⁴

$$\tau_i^n = \inf \left\{ t > \tau_{i-1}^n : \left| z_t - z_{\tau_{i-1}^n} \right| \geq 2^{-n}, t \in \mathbb{R}_+ \right\} \quad \text{and} \quad \tau_0^n \equiv 0.$$

Define $I^n : D(\mathbb{R}_+, \mathbb{R}) \times D(\mathbb{R}_+, \mathbb{R}) \rightarrow D(\mathbb{R}_+, \mathbb{R})$ as

$$(5.1) \quad I^n(z, x)(t) = z(0)x(0) + \sum_{i \geq 0} z(\tau_i^n) (x(\tau_{i+1}^n \wedge t) - x(\tau_i^n \wedge t))$$

and set $I(z, x) = \lim_n I^n(z, x)$ whenever this limit exists in the topology of uniform convergence on compact sets on $D(\mathbb{R}_+, \mathbb{R})$, otherwise set $I(z, x) = 0$.

¹⁴With the convention $\inf \emptyset = \infty$ and $\tau_{i+1}^n = \infty$ if $\tau_i^n = \infty$.

Proof of Theorem 33. We define

$$(5.2) \quad J_Z^i(t, \omega) := \begin{cases} I(Z(\omega), X^i(\omega))_t & , \text{if } t < \zeta(\omega) \\ 0 & , \text{else} \end{cases}$$

where $I(\cdot, \cdot) : D(\mathbb{R}_+, \mathbb{R}) \times D(\mathbb{R}_+, \mathbb{R}) \rightarrow D(\mathbb{R}_+, \mathbb{R})$ is the map constructed in Theorem 35 and Remark 38. For brevity write J instead of J_Z^i throughout this proof. Point (2) follows immediately from Theorem 35.

To see the causal differentiability, consider $(t, \omega) \in \mathbb{R}_+ \times \Omega$ with $t < \zeta(\omega)$. In this case, if the adapted Riemann sum (5.1) of the Bichteler–Karandikar integral $I(Z(\omega), X^i(\omega))(t)$ converges as a uniform limit then this is also the case for $\omega^{t,r}$ or $\omega_{\wedge t}$ for any $r = (r^1, \dots, r^d) \in \mathbb{R}^d$: first note that Z is \mathcal{O}^0 -optional (since \mathcal{O}^0 is generated by the càdlàg processes) and so by Proposition 1 the trajectories of $Z(\omega)$ and $Z(\omega_{\wedge t})$ coincide on $[0, t]$, hence the sequence $(\tau_i^n(\omega))_{i \geq 0}$ is the same as $(\tau_i^n(\omega_{\wedge t}))_{i \geq 0}$ when restricted to $[0, t]$. Then directly from the definition of the Bichteler–Karandikar integral as limit of (5.1) it follows that

$$J(t + \epsilon, \omega_{\wedge t}) = I(Z(\omega_{\wedge t}), X^i(\omega_{\wedge t}))_{t+\epsilon} = I(Z(\omega_{\wedge t}), X^i(\omega))_t = I(Z(\omega), X^i(\omega))_t,$$

which implies $\partial_0 J(t, \omega) = 0$. Similarly, we can use that the trajectories of $Z(\omega^{t,r})$ and $Z(\omega)$ coincide on $[0, t]$ to see that

$$\begin{aligned} J(t, \omega^{t,r}) &= I(Z(\omega^{t,r}), X^i(\omega^{t,r}))_t = I(Z(\omega^{t,r}), X^i(\omega))_t + Z_{t-}(\omega^{t,r}) r^i \\ &= I(Z(\omega), X^i(\omega))_t + Z_{t-}(\omega) r^i \end{aligned}$$

which shows the claimed first space derivative. By \mathcal{O}^0 -measurability (resp. causality) of Z

$$Z_{t-}(\omega^{t,\Delta}) = \lim_{r \nearrow t} Z_r(\omega^{t,\Delta}) = \lim_{r \nearrow t} Z_r((\omega^{t,\Delta})_{\wedge r}) = Z_{t-}(\omega^{t,-\Delta t \omega})$$

and similarly $Z_{t-}(\omega) = Z_{t-}(\omega^{t,-\Delta t \omega})$ which implies $\Delta J_t(\omega) = 0$.

To see that $J \in C^{1,2}$ we first show that J is a continuous, regular functional, more precisely that $\forall \mathbb{P} \in \mathcal{M}_c^{semi}$ the map $(t, \omega) \mapsto J_t(\omega^n)$ is an adapted càdlàg process that converges ucp to the (\mathbb{P} -a.s.) continuous process $(t, \omega) \mapsto J_t(\omega)$ on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$. Therefore fix $\mathbb{P} \in \mathcal{M}_c^{semi}$, any sequence of partitions $(\pi(n))_n$ converging to the identity and denote the associated piecewise constant approximation of ω with $(\omega^n)_n$. From the very definition of J we have $\forall (t, \omega) \in \mathbb{R}_+ \times \Omega$

$$J_t^n(\omega) := J_t(\omega^n) = I(Z(\omega^n), X^i(\omega^n))(t).$$

Now,

$$(t, \omega) \mapsto Z_t^n(\omega) := Z_t(\omega^n) \text{ and } (t, \omega) \mapsto X_t^n(\omega) := X_t^i(\omega^n)$$

are càdlàg processes on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$ and X^n is of bounded variation (on compacts) and therefore a semimartingale, hence Theorem 35 guarantees that

$$J^n(\omega) = \left(\int_0^\cdot Z_-^n dX^n \right)(\omega) \text{ for } \mathbb{P} - a.e. \omega.$$

so especially J^n is a càdlàg process. To see that $J_-^n \rightarrow J_-$ ucp note that for every n we have \mathbb{P} -a.s.

$$\begin{aligned} \int_0^\cdot Z_-^n dX^n &= \int_0^\cdot (Z_-^n - Z_-) dX^n + \int_0^\cdot Z_- dX^n \\ &= \sum_{k: t_k^n \leq \cdot} (Z_-^n - Z_-)_{t_{k-1}^n} X_{t_{k-1}^n, t_k^n} + \sum_{k: t_k^n \leq \cdot} Z_{t_{k-1}^n} - X_{t_{k-1}^n, t_k^n} \\ &= \int_0^\cdot C^n dX + \sum_{k: t_k^n \leq \cdot} Z_{t_{k-1}^n} - X_{t_{k-1}^n, t_k^n} \end{aligned}$$

with $C_t^n = \sum_{\pi(n)} (Z_{t_k^n}^n - Z_{t_{k-1}^n}^n) 1_{[t_k^n \wedge \cdot, t_{k+1}^n \wedge \cdot)}$ (the second equality holds since X^n is of bounded variation on compacts, hence the stochastic integrals are indistinguishable from pathwise Lebesgue integrals, cf. [32, Theorem 18]). Since Z is a regular functional, $C^n \rightarrow 0$ ucp and the first integral converges to 0 ucp as $n \rightarrow \infty$. Standard results show ucp convergence of the Riemann sums to $\int_0^\cdot Z_- dX$. \square

As an immediate but important corollary we get the existence of a $C^{1,2}$ version of the quadratic variation $[X]$ of X .

Corollary 39. *For every $i, j \in \{1, \dots, d\}$ there exists a causal functional*

$$B^{ij} : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$$

such that

- (1) $B^{ij} \in C^{1,2}$,
- (2) B^{ij} is indistinguishable from $[X^i, X^j]$ on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P})$ for every $\mathbb{P} \in \mathcal{M}_c^{semi}$, i.e.

$$B^{ij} = [X^i, X^j] \quad \mathbb{P} \text{ -a.s.}$$

Moreover $\forall (t, \omega) \in \mathbb{R}_+ \times \Omega, t < \zeta(\omega), i, j \in \{1, \dots, d\}$

$$\partial_0 B_t^{ij}(\omega) = 0$$

and

$$\begin{aligned} \nabla B_t^{ij}(\omega) &= (\Delta_t X^j(\omega) 1_{k=i} + \Delta_t X^i(\omega) 1_{k=j})_{k=1, \dots, d} \\ \Delta B_t^{ij}(\omega) &= (1_{l=j} 1_{k=i} + 1_{l=i} 1_{k=j})_{k, l=1, \dots, d} \end{aligned}$$

We also use the notation $[X^i, X^j]^{BK}$ for B^{ij} .

Proof. Since X^i is the coordinate process it fulfills the assumptions for the integrand in Theorem 33 except that X^i is not real-valued (only \mathbb{P} -a.s.). Therefore set $\overline{X}_t(\omega) = X_t(\omega) 1_{\zeta(\omega)=\infty}$. Hence, Theorem 33 allows us to define

$$(5.3) \quad B_t^{ij}(\omega) = \overline{X}_t^i(\omega) \overline{X}_t^j(\omega) - \int_0^t \overline{X}_-^i(\omega) d^{BK} X^j(\omega) - \int_0^t \overline{X}_-^j(\omega) d^{BK} X^i(\omega)$$

Points (1) & (2) follow from Theorem 35, Theorem 33 and the identity

$$[X^i, X^j] = X^i X^j - \int_0^\cdot X^i dX^j - \int_0^\cdot X^j dX^i.$$

Finally, apply the causal derivatives to $B^{ij}(t, \omega)$ which gives for the time derivative

$$\partial_0 (X^i X^j)_t(\omega) = \partial_0 \left(\int_0^\cdot X^i(\omega) d^{BK} X^j(\omega) \right)_t = \partial_0 \left(\int_0^\cdot X^j(\omega) d^{BK} X^i(\omega) \right)_t = 0.$$

and from Theorem 33 the space derivatives of the integrals follows as well

$$\partial_i B_t^{ij}(\omega) = \overline{X}_t^j(\omega) - \overline{X}_{t-}^j(\omega) = \Delta_t \overline{X}^j(\omega).$$

□

Remark 40. The first spatial derivative of B^{ij} is indistinguishable under every continuous semimartingale measure from 0 but the second spatial derivative is not! This is a direct consequence of the pathwise nature of the causal derivatives as pointed out in Remark 8 (pathwise the first derivative is not equal to 0).

Remark 41. It would be desirable to have a result which guarantees the existence of a $C^{1,2}$ -aggregator, i.e. given a reasonably nice subset \mathcal{P} of \mathcal{M}_c^{semi} and a family of processes

$$\left\{ Z^\mathbb{P} : \mathbb{P} \in \mathcal{P}, Z^\mathbb{P} \text{ is a progr. measurable process on } (\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P}) \right\}$$

which fulfill some consistency condition then there exists a causal functional $Z \in C^{1,2}$ s.t. $Z = Z^\mathbb{P}$ \mathbb{P} -a.s. $\forall \mathbb{P} \in \mathcal{P}$. Indeed, the only result we are aware of in this direction is the main result in [33, Theorem 5.1] which shows for an important subset \mathcal{P} of \mathcal{M}_c^{semi} the existence of an aggregator. However, even with the restriction to \mathcal{P} it is not clear which conditions would guarantee that the aggregator is causally differentiable which is why we have chosen a constructive approach for the two examples of stochastic integration and quadratic variation which also applies to all elements of \mathcal{M}_c^{semi} .

We give some applications of Theorem 33 and Corollary 39: the first is an easy application of the chain and product rule.

Example 42. Let $d = 1$ and $E_t(\omega) = \exp\left(X_t(\omega) - \frac{1}{2}[X]_t^{BK}(\omega)\right)$. Then $\partial_0 E_t(\omega) = 0, \nabla E_t(\omega) = E_t(\omega)(1 - \Delta_t X(\omega))$ and $\Delta E_t(\omega) = E_t(\omega)\Delta_t X(\omega)(-2 + \Delta_t X(\omega))$. Hence, Theorem 22 reduces to the classic identity

$$E_\cdot = 1 + \int_0^\cdot E_r dX \text{ on } (\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P}), \mathbb{P} \in \mathcal{M}_c^{semi}.$$

The path-dependent Itô-formula, Theorem 22, shows that every element of $C^{1,2}$ is a Itô-process \mathcal{M}_c^{semi} -quasi-sure. Using Theorem 35 we can now go the other direction:

Example 43. Assume $\mu, \sigma : \mathbb{R} \times \Omega \rightarrow \mathbb{R}$ fulfill the same assumptions as Z in Theorem 33. Then the Itô-process

$$\int_0^\cdot \mu_r dr + \int_0^\cdot \sigma_{r-} \cdot dX_r = \int_0^\cdot \mu_r dr + \sum_{i=1}^d \left(\int_0^\cdot \sigma_{r-}^i dX_r^i \right)$$

is indistinguishable on $(\Omega, \mathcal{F}^\mathbb{P}, \mathcal{F}_t^\mathbb{P}, \mathbb{P}), \mathbb{P} \in \mathcal{M}_c^{semi}$, from $F \in C^{1,2}$ defined as

$$F_t(\omega) := \int_0^t \mu_r(\omega) dr + \sum_{i=1}^d \int_0^t \sigma_{r-}^i(\omega) d^{BK} X_r^i(\omega)$$

and $\forall (t, \omega) \in \mathbb{R}_+ \times \Omega, t < \zeta(\omega)$,

$$\partial_0 F_t(\omega) = \mu_t(\omega), \nabla F_t(\omega) = \sigma_{t-}(\omega) = \left(\sigma_{t-}^d(\omega), \dots, \sigma_{t-}^1(\omega) \right)^T, \Delta F_t(\omega) = 0.$$

(hence $\partial_0 F = \mu, \nabla F = \sigma_-$ \mathbb{P} -a.s.) which captures the intuition that the causal time derivative measures the infinitesimal drift and the causal space derivative measures sensitivity to instantaneous changes of the underlying process X , cf. Remark 23.

Even if we are only interested in a fixed measure $\mathbb{P} \in \mathcal{M}_c^{semi}$, Theorem 22 and the results of this section are a non-trivial extension of the functional Itô-formula: while not the topic of this article, we finish with an application of Corollary 39 to mathematical finance which allows to compare it with the approach of Cont–Fournie [9, 8] who treat the same example (see [20, Section 5] for more applications in finance relying on such a generalized functional Itô-formula).

Example 44. For fixed $\mathbb{P} \in \mathcal{M}_c^{semi}$ the coordinate process $X = (X^i)_{i=1}^d$ is interpreted as the discounted price process of d assets in a security market (for brevity we let $d = 1$ and assume no interest rates). Assume there exists an equivalent martingale measure $\mathbb{Q} \in \mathcal{M}_c^{loc}, \mathbb{Q} \sim \mathbb{P}$ and $[X] = \int \sigma_r^2 dr$ under \mathbb{Q} for some adapted, continuous, bounded process σ^2 . The usual Black–Scholes–Merton dynamic replication argument translates and shows that if $F \in C^{1,2}$ and

$$\partial_0 F + \frac{\sigma^2}{2} \Delta F = 0, F_T = \pi \mathbb{P} - a.s.$$

then F_t is an arbitrage free price for the claim π at time $t \in [0, T]$ (i.e. a version of $\mathbb{E}_\mathbb{Q}[\pi | \mathcal{F}_t^\mathbb{P}]$)¹⁵. Applied to variance swaps we can make the Ansatz $\pi = [X]_T^{BK} - k$ and $F_t(\omega) = [X]_t^{BK}(\omega) - k + v(t, X_t(\omega))$ and above reduces to the parabolic SPDE

$$\frac{dv}{dt} + \frac{\sigma_t^2}{2} \frac{d^2 v}{dx^2} = -\sigma_t^2, v(T, X_T) = 0 \mathbb{P} - a.s.$$

(for a local vol. model $\sigma_t^2 = \sigma_0^2(t, X_t) X_t^2$ this further reduces to the backward heat equation with source term $x^2 \sigma_0(t, x)$). Of course, more interesting applications exist for more complex payoffs (nonlinear options on variance, etc.) in combination with the uncertainty $\mathbb{P} \in \mathcal{P}$ for $\mathcal{P} \subset \mathcal{M}_c^{semi}$ but the related pathdependent PDEs and questions of uniqueness, existence, interplay with the nonlinear expectation $\mathcal{E}(\cdot) = \inf_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_\mathbb{P}[\cdot]$ etc. touch the tip of an iceberg which is currently under heavy development even for the case when π is a continuous functional in uniform norm of the underlying — we draw attention to work of Peng–Wang [30, 31] and work of Ekren–Keller–Touzi–Zhang [13, 14, 15].

¹⁵If $\sigma_t^2 = \sigma_{BS}^2 X_t^2, \sigma_{BS}^2 \in \mathbb{R}$ fix and $\pi = (X_T - K)^+$ one makes the Ansatz $F_t = f(t, X_t)$ and this reduces to the classic Black–Scholes–Merton PDE. The usual *Greeks* $\partial_0 F, \nabla F, \Delta F$ are now functionals and it is interesting to recall Proposition 29 and Remark 31 which show that as a process, only $\partial_0 F + \frac{\sigma^2}{2} \Delta F$ is uniquely defined, i.e. in a financial context it makes sense to think of this non-uniqueness as the Gamma–Theta or convexity–time-decay tradeoff.

Acknowledgement 45. HO is grateful for support from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement nr. 258237 and DFG Grant SPP-1324. Further, the author would like to thank Rama Cont for helpful remarks.

REFERENCES

- [1] Hyungsok Ahn. Semimartingale integral representation. *Ann. Probab.*, 25(2):997–1010, 1997.
- [2] R. Azencott. Grandes deviations et applications. *Ecole d'été de probabilités de Saint-Flour VIII-1978*, pages 1–176, 1980.
- [3] J. Bertoin. *Lévy processes*, volume 121. Cambridge Univ Pr, 1998.
- [4] K. Bichteler. Stochastic integration and L^p -theory of semimartingales. *The Annals of Probability*, 9(1):49–89, 1981.
- [5] M.W. Botsko and R.A. Gossler. Stronger versions of the fundamental theorem of calculus. *American Mathematical Monthly*, pages 294–296, 1986.
- [6] R. Buckdahn and J. Ma. Pathwise stochastic taylor expansions and stochastic viscosity solutions for fully nonlinear stochastic pdes. *The annals of Probability*, 30(3):1131–1171, 2002.
- [7] J.M.C. Clark. The design of robust approximations to stochastic differential equations in nonlinear filtering. In J.K. Skwirzynsky, editor, *Communication Systems in Random Processes Theory*. Sijthoff, Nordhoff, 1978.
- [8] R. Cont and D. Fournie. Change of variable formulas for non-anticipative functionals on path space. *Journal of Functional Analysis*, 259:1043–1072, 2010.
- [9] R. Cont and D. Fournie. Functional ito calculus and stochastic integral representation of martingales. *Annals of Probability*, 2011. In press.
- [10] Claude Dellacherie and Paul-André Meyer. *Probabilities and potential*, volume 29 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1978.
- [11] L. Denis and C. Martini. A theoretical framework for the pricing of contingent claims in the presence of model uncertainty. *The Annals of Applied Probability*, pages 827–852, 2006.
- [12] B. Dupire. Functional Ito calculus. 2010. SSRN.
- [13] I. Ekren, C. Keller, N. Touzi, and J. Zhang. On Viscosity Solutions of Path Dependent PDEs. *ArXiv e-prints*, September 2011.
- [14] I. Ekren, N. Touzi, and J. Zhang. Viscosity Solutions of Fully Nonlinear Parabolic Path Dependent PDEs: Part I. *ArXiv e-prints*, September 2012.
- [15] I. Ekren, N. Touzi, and J. Zhang. Viscosity Solutions of Fully Nonlinear Parabolic Path Dependent PDEs: Part II. *ArXiv e-prints*, September 2012.
- [16] Wendell H. Fleming and H. Mete Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Stochastic Modelling and Applied Probability*. Springer, New York, second edition, 2006.
- [17] M. Fliess. Fonctionnelles causales non linéaires et indéterminées non commutatives. *Bull. Soc. Math. France*, 109(1):3–40, 1981.
- [18] M. Fliess. On the concept of derivatives and taylor expansions for nonlinear input-output systems. In *Decision and Control, 1983. The 22nd IEEE Conference on*, volume 22, pages 643–646. IEEE, 1983.
- [19] H. Föllmer. Calcul d'Itô sans probabilités. In *Seminar on Probability, XV (Univ. Strasbourg, Strasbourg, 1979/1980) (French)*, pages 143–150. Springer, Berlin, 1981.
- [20] David Fournie. *Functional Ito calculus and applications*. PhD thesis, Columbia University, 2010.
- [21] M.I. Freidlin and A.D. Wentzell. *Random perturbations of dynamical systems*, volume 260. Springer Verlag, 1998.
- [22] Peter Friz and Harald Oberhauser. Rough path limits of the Wong-Zakai type with a modified drift term. *J. Funct. Anal.*, 256:3236–3256, 2009.
- [23] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*. North-Holland Publishing Co., Amsterdam, second edition, 1989.
- [24] Jean Jacod and Albert N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, second edition, 2003.
- [25] R. L. Karandikar. On pathwise stochastic integration. *Stochastic Processes and their Applications*, 57(1):11 – 18, 1995.
- [26] S. Kusuoka and D.W. Stroock. Precise asymptotics of certain wiener functionals. *Journal of functional analysis*, 99(1):1–74, 1991.
- [27] E. J. McShane. Stochastic differential equations and models of random processes. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971)*, Vol. III: *Probability theory*, pages 263–294. Berkeley, Calif., 1972. Univ. California Press.
- [28] S. Peng. G-Brownian Motion and Dynamic Risk Measure under Volatility Uncertainty. *ArXiv e-prints*, November 2007.
- [29] S. Peng. Nonlinear expectations and stochastic calculus under uncertainty. *arXiv preprint arXiv:1002.4546*, 2010.
- [30] S. Peng. Note on Viscosity Solution of Path-Dependent PDE and G-Martingales. *ArXiv e-prints*, June 2011.
- [31] S. Peng and F. Wang. BSDE, Path-dependent PDE and Nonlinear Feynman-Kac Formula. *ArXiv e-prints*, August 2011.
- [32] Philip E. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, second edition, 2004. Stochastic Modelling and Applied Probability.
- [33] M. Soner, N. Touzi, and J. Zhang. Quasi-sure stochastic analysis through aggregation. *Electronic Journal of Probability*, 16:1844–1879, 2011.

- [34] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1979.
- [35] Héctor J. Sussmann. Limits of the Wong-Zakai type with a modified drift term. In *Stochastic analysis*, pages 475–493. Academic Press, Boston, MA, 1991.
- [36] S.R.S. Varadhan and SRS Varadhan. *Large deviations and applications*, volume 46. SIAM, 1984.
- [37] E. Wong and M. Zakai. On the convergence of ordinary integrals to stochastic integrals. *The Annals of Mathematical Statistics*, pages 1560–1564, 1965.
- [38] E. Wong and M. Zakai. On the relation between ordinary and stochastic differential equations. *International Journal of Engineering Science*, 3(2):213–229, 1965.

TU BERLIN, INSTITUT FÜR MATHEMATIK, STRASSE DES 17. JUNI 136, 10623 BERLIN
E-mail address: `h.oberhauser@gmail.com`